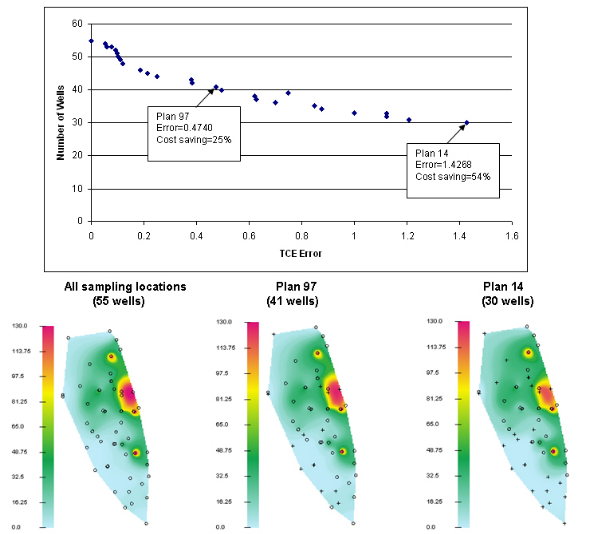
ESTCP
Cost and Performance Report

(ER-0629)



Adaptive Long-Term Monitoring at Environmental Restoration Sites

November 2009

ENVIRONMENTAL SECURITY
TECHNOLOGY CERTIFICATION PROGRAM

U.S. Department of Defense

Report Documentation Page	*Form Approved* *OMB No. 0704-0188*

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **NOV 2009**	2. REPORT TYPE	3. DATES COVERED **00-00-2009 to 00-00-2009**

4. TITLE AND SUBTITLE **Adaptive Long-Term Monitoring at Environmental Restoration Sites**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Environmental Security Technology Certification Program (ESTCP),4800 Mark Center Drive, Suite 17D08,Alexandria,VA,22350-3605**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**59**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

COST & PERFORMANCE REPORT
Project: ER-0629

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (continued)

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

ACRONYMS AND ABBREVIATIONS

BTEX	benzene, toluene, ethylbenzene, and xylene
c12DCE	cis-1,2-dichloroethene
COC	contaminant of concern
CSV	comma separated variable
D2K	Data to Knowledge
DoD	Department of Defense
DT	Data Tracker
EM CX	Environmental and Munitions Center of Expertise (formerly the Hazardous, Toxic, and Radioactive Waste Center of Expertise)
EMO	Evolutionary Multi-objective Optimizer
ESTCP	Environmental Security Technology Certification Program
GA	genetic algorithm
GAFB	former George Air Force Base
HMSI	Hazard Management Systems, Inc.
IDW	inverse distance weighting
LTM	long term monitoring
LTMO	long-term monitoring optimization
MAROS	Monitoring and Remediation Optimization Software
NAVFAC ESC	Naval Facilities Engineering Command/Engineering Service Center
NCSA	National Center for Supercomputing Applications
NSGA-II	Nondominated Sorted Genetic Algorithm-II
NOP	former Nebraska Ordnance Plant
OMB	Office of Management and Budget
P&T	pump-and-treat
QA/QC	quality assurance/quality control
RDX	Royal Demolition Explosive, or Hexahydro-1,3,5-trinitro-1,3,5-triazine
SO	Sampling Optimizer
STO	spatiotemporal optimization
TCE	trichloroethene
TRECC	Technology, Research, Education, and Commercialization Center

ACRONYMS AND ABBREVIATIONS (continued)

USACE	U.S. Army Corps of Engineers
USEPA	U.S. Environmental Protection Agency
VC	vinyl chloride

ACKNOWLEDGEMENTS

The project team consisted of the following organizations:

- Karla Harre and Tanwir Chaudhry (Naval Facilities Engineering Command/Engineering Service Center [NAVFAC ESC])
- Rob Greenwald, Weiwei Jian, and Yan Zhang (GeoTrans, Inc. [GeoTrans])
- Matt Zavislak, John Dustman, and Dr. Barbara Minsker (Consultant) (Summit Envirosolutions, Inc. [Summit])
- Dr. Charles Davis (Environmetrics & Statistics Limited [EnviroStat])

NAVFAC ESC led the project, including management of project funds, schedule, and deliverables. GeoTrans was responsible for conducting demonstration analysis for selected sites, coordinating preparation of reports, and coordinating site visits. Summit is the developer of the software and offered training, technical support, and assistance regarding preparation of the reports. EnviroStat prepared data for optimization analysis as well as artificial anomalies for testing of Data Tracker, provided technical support regarding validation of the results, and participated in report preparation and review.

In addition, we would like to acknowledge the participation of the following individuals and/or organizations:

- Dave Becker, U.S. Army Corps of Engineers (USACE), Environmental and Munitions Center of Expertise (EM CX) (Omaha)—Mr. Becker participated in numerous conference calls, provided technical support and review of reports, and coordinated the participation of one of the demonstration sites.
- Dave Wilson, U.S. Environmental Protection Agency (USEPA) Region V—Mr. Wilson coordinated an additional application of the software by a USEPA contractor at a site in Region V and provided feedback regarding that effort to our project team.
- Mindy Vanderford, GSI Environmental, Inc.—Ms. Vanderford provided guidance and feedback regarding the application of the Monitoring and Remediation Optimization Software (MAROS) at one of the demonstration sites.

We would also like to thank all site personnel and contractors at the three demonstration sites associated with this Environmental Security Technology Certification Program (ESCTP) project. The demonstration sites were as follows:

- Former George Air Force Base (GAFB) Site, Victorville, CA
- Former Nebraska Ordnance Plant (NOP) Site, Mead, NE
- Camp Allen Landfill Site, Norfolk, VA

Their time and efforts are greatly appreciated.

Technical material contained in this report has been approved for public release.

This page left blank intentionally.

1.0 EXECUTIVE SUMMARY

The objective of this ESTCP project was to demonstrate and validate the use of the Sampling Optimizer (SO) and Data Tracker (DT) software (the Summit Software), offered by Summit, at three Department of Defense (DoD) sites. The three demonstration sites were as follows:

- Former George Air Force Base (GAFB) Site, Victorville, CA
- Former Nebraska Ordnance Plant (NOP) Site, Mead, NE
- Camp Allen Landfill Site, Norfolk, VA

Monitoring and Remediation Optimization Software (MAROS) (developed by GSI Environmental, Inc.) was also applied at one of the three demonstration sites. The Summit Software demonstrated in this ESTCP project provides a set of tools for long-term monitoring optimization (LTMO) and consists of two major modules:

- SO identifies redundant sampling locations (spatial optimization), or redundant locations and frequencies (spatiotemporal optimization), in historical data.

- DT allows current monitoring data to be reviewed against selected historical data (i.e., the background data) to identify cases where current data deviate from expectations that are based on the background values and patterns.

Model Builder is an additional component within the software with two functions—one for model fitting, visualization, and analysis (with kriging or inverse distance weighting) and another for visualizing relative uncertainty.

Key results of the project include the following:

- The software is easy to learn and use, and no bugs or software errors are apparent.

- Of the six options available, kriging with quantile data transformation qualitatively provided the best representation of the plumes with Model Builder.

- SO provided useful trade-off curves of sampling cost versus the interpolation error that resulted from removing samples.

- Potential savings from spatial analysis (i.e., eliminating redundant sampling locations), calculated based on number of wells eliminated, ranged from approximately 10% to 67%. A value of approximately 35% appears to be representative. For the two sites where spatiotemporal analysis was performed, the potential savings, based on number of samples eliminated per year, only ranged from approximately 4% to approximately 17%. Reasons that overly conservative results may be obtained from the spatiotemporal analysis (versus spatial analysis) are discussed in the Final Report.

- The software allows redundancy analysis to be performed simultaneously for multiple contaminants of concern. Based on the testing performed, it appears that simultaneous evaluation is preferable to evaluating multiple constituents independently.

- Validation exercises conducted with reserved data provided confidence in the results provided by SO at all three demonstration sites.

- If distribution of sampling locations is not consistent between events, the results of relative mass or mass flux calculations provided by the software may be questionable, for reasons discussed in the Final Report. This limitation also exists for MAROS.

- The relative uncertainty maps provided by Model Builder were not found to be particularly useful.

- DT identified as "out-of-bounds" the vast majority of artificial anomalies added by EnviroStat for testing of this module and also identified some anomalies in the actual site data.

- Some limitations of the software were identified that could potentially be mitigated by future software improvements, including the following:

    - The software interpolates spatially but does not perform interpolations in time. The software would be improved if there was a feature to optionally fill in missing values via temporal interpolation.

    - In DT, the software would be improved if plots of concentration versus time used different symbols to differentiate between the background data and the current data.

    - In DT, the software would be improved if selected values could be imported with a flag so that they can be included on concentration versus time plots (with a different symbol) but not used to calculate the prediction limits.

    - The DT portion of the software would be improved if it also indicated whether the concentration trend for a specific contaminant of concern (COC) at a specific well is increasing, stable, or decreasing (i.e., as is provided by MAROS).

    - The software import features would be improved by adding features to consolidate data into discrete events and to recognize flags (e.g., for nondetect values).

The level of effort and computation time for applying the software at the demonstration sites, and a basis for estimating the costs of applying the software at other sites, are provided in this report. The software and Users Guide, additional components of this ESTCP project, are now available for use at government sites by government personnel and their contractors.

This page left blank intentionally.

2.0 INTRODUCTION

2.1 BACKGROUND

LTM provides a mechanism for evaluating performance of groundwater remedies and is essential for ensuring protection of human health and the environment. The costs of future monitoring are expected to be substantial since LTM generally spans many years and is required at a large number of sites. Efficiency of LTM can be improved by the following:

- Obtaining only the essential data needed for monitoring current conditions by eliminating redundant sampling locations and/or frequencies

- Using a semi-automated approach to identify values from recently collected data that are not within expectations (based on statistical evaluation of previous values)

- Tracking performance relative to specific metrics (e.g., assessing reductions in overall contaminant mass).

This project demonstrates the application of the SO and DT software offered by Summit, which is intended to address the items listed above. We sometimes refer to this software as the Summit Software.

The evaluation of data redundancy in the SO portion of the software uses mathematical optimization, which is unique relative to other LTMO software products. This allows sampling redundancy to be evaluated on a system-wide basis (e.g., best solution if one location is removed, if two locations are removed, if three locations are removed, etc.). A key benefit of this approach is that it allows the trade-off between the number of samples and the accuracy of the resulting plume interpolation to be assessed. This is a significant improvement over the approach for evaluating data redundancy utilized in the MAROS (developed by GSI Environmental Inc.), which is not based on mathematical optimization. In MAROS, individual wells locations are evaluated for redundancy based on impacts of removing that well alone; consequently, the impact of removing groups of wells cannot be assessed and the aforementioned trade-off cannot be evaluated. Another key benefit of the Summit approach for evaluating data redundancy is that plume visualizations for the baseline plan (i.e., all samples) versus improved plans (i.e., reduced numbers of samples) are created within the software. These comparative visualizations are quite effective for communication with stakeholders and regulators.

The DT portion of the software identifies values from recently collected data that are not within expectations (based on statistical evaluation of previous values). This is somewhat different in approach and implementation versus MAROS, which evaluates concentration trends over time at individual wells as increasing, decreasing, or stable. Rather than indicate increasing versus decreasing trends, the intent of the DT portion of the software is to automatically highlight which recently collected data values are unexpected and require further attention. These unexpected values may be due to significant increasing or decreasing trends, or may be due to "bad data" (lab error, sampling error, database error, etc.). DT allows for "expected" time trends to be

either stable or smoothly decreasing; the latter is appropriate for monitoring an effective passive remediation system, for example.

2.2 OBJECTIVE OF THE PROJECT

The objective of this project is to demonstrate and validate the use of the SO and DT software (i.e., the Summit Software) by applying the software at three DoD sites. A secondary objective is to compare the results with MAROS at one site. Another component of the project is to transfer the software and documentation to the government for free use at government sites by government personnel and their contractors.

2.3 REGULATORY DRIVERS

There are no regulatory issues directly associated with this effort, although there has been a general focus in recent years regarding optimization of all facets of remediation including LTM. The software demonstrated in this project is intended to improve the efficiency and assessment of the monitoring well networks and data that are collected in current monitoring events, which will ultimately address regulatory objectives and allow for improved communication between all site stakeholders. Implementation of revised sampling plans suggested for the demonstration sites is not within the scope of this project.

3.0 TECHNOLOGY

3.1 TECHNOLOGY DESCRIPTION

The Summit Software is a set of desktop software tools consisting of two major modules, SO and DT.

- *Sampling Optimizer* identifies redundant sampling locations (spatial optimization), or redundant locations and frequencies (spatiotemporal optimization), in historical data. This module identifies redundancies using a multi-objective genetic algorithm (GA) to obtain monitoring designs that represent optimal tradeoffs among two or more monitoring objectives, such as minimizing the number of samples and minimizing the interpolation error at locations that are removed. The error will generally increase as number of wells decreases, resulting in a trade-off.

- *Data Tracker* allows current monitoring data to be reviewed against selected historical data (i.e., the background data) to identify cases where current data deviate from expectations that are based on the background values and patterns.

Model Builder is an additional component within the software that is utilized by SO and, in some cases, by DT. Model Builder has two sections: one for model fitting, visualization, and analysis (with kriging or inverse distance weighting) and another for visualizing relative uncertainty.

The software modules/components listed above (SO, DT, and Model Builder) are highlighted in Figure 1, which is a general flowchart illustrating the application of the software. The software is applied in several ways:

- During initial optimization and periodic review (upper portion of the figure), Model Builder constructs spatial and/or spatiotemporal models for the measurements of the primary COCs. The model identified by Model Builder is then used by SO to identify optimal sampling plans for subsequent routine monitoring (i.e., with redundancies eliminated).

- After new sampling events (lower portion of the figure), the software can be used during routine monitoring to identify anomalies or departures from expectations in the primary COCs and other COCs as desired. In addition, changes over time in plume mass and/or mass flux across a boundary can be tracked. If the quantity of interest involves modeling for each monitoring event (e.g., contaminant mass based on interpolation), that modeling is provided by Model Builder.

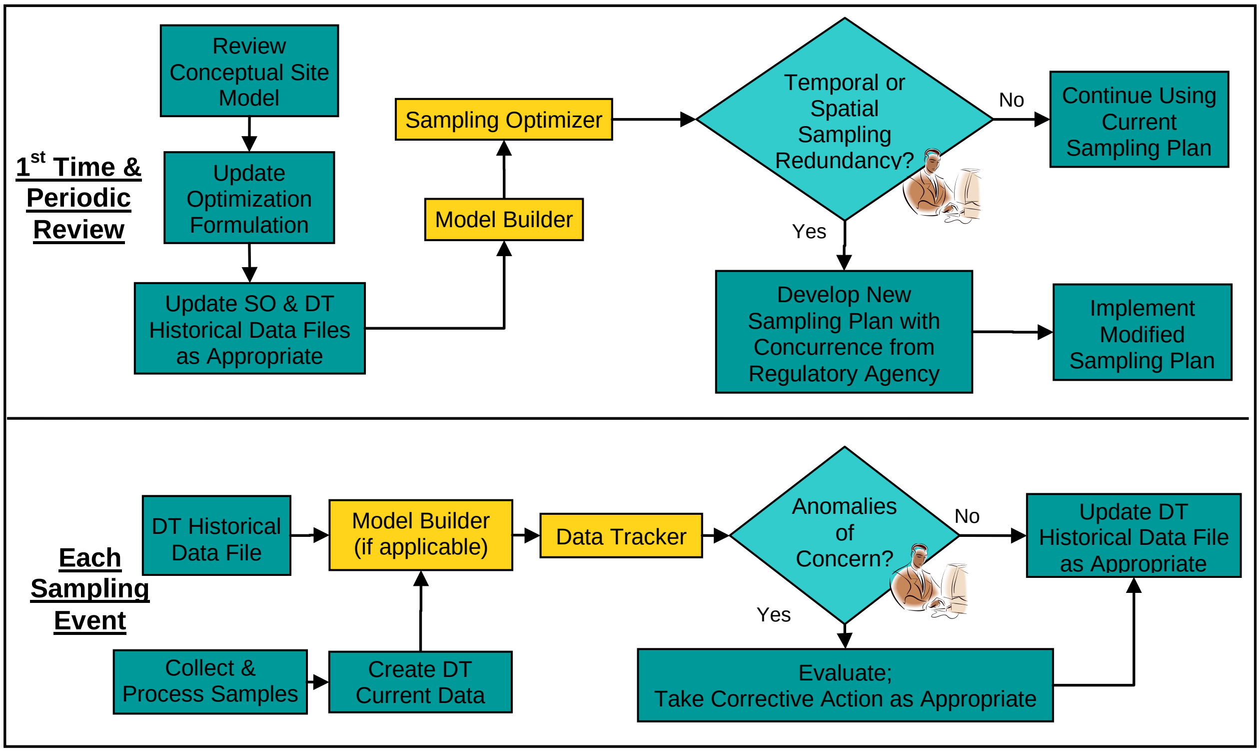


Figure 1. General flowchart of software application.

Although the software provides specific results as output, there is still an aspect of interpretation required by an analyst, as illustrated in Figure 1. With respect to the data redundancy evaluation, the software produces an optimal trade-off curve (example provided on Figure 2), and the analyst must choose specific sampling plans along the trade-off curve for further evaluation. For each number of wells, SO evaluates many potential sampling plans; the sampling plans with the least errors are those shown in the trade-off curves.

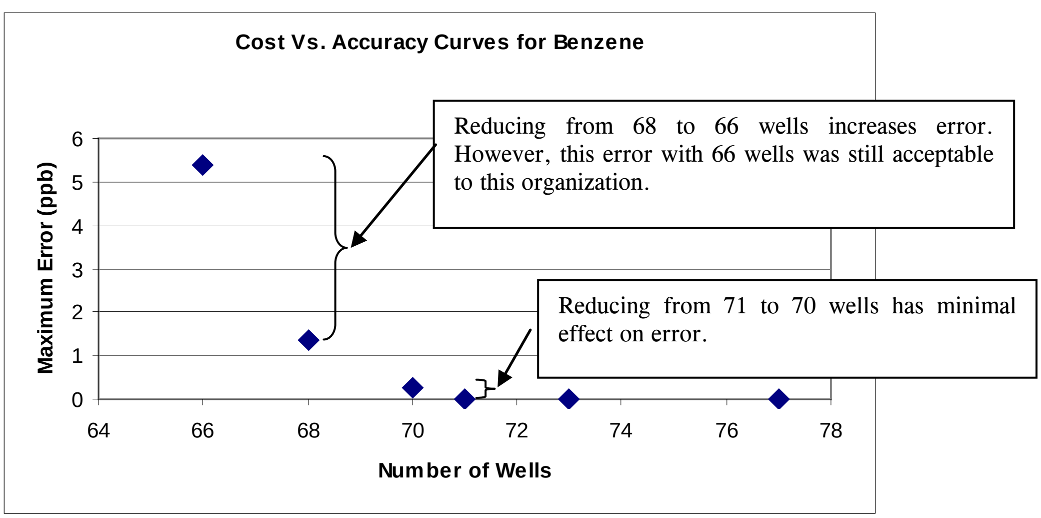


Figure 2. Example of a trade-off curve for spatial optimization.

The Summit Software calculates the error as a dimensionless parameter, using a tool called the Cutoff Error Calculator, which is described in detail in the Final Report. The error is an indicator of the maximum difference between an interpolated value and an actual value at any location where a sample is eliminated.

With respect to the detection of anomalies with DT, the software identifies values in recently collected data that are "out-of-bounds" and provides a graph of concentration versus time for visual review. The analyst must then determine if any response or action is appropriate, such as correction of an erroneous laboratory report or further investigation of a potential new source of contamination. Thus, the corrective action or response referred to in the figure can take many forms depending on the nature of the anomaly.

The Summit tools were built on technology development and research at the University of Illinois from 1997 to 2004. Reed et al. (2003), with funding from an U.S. Environmental Protection Agency (USEPA) STAR Fellowship, created an automated methodology for setting parameter values for the Nondominated Sorted Genetic Algorithm-II (NSGA-II). A Technology, Research, Education, and Commercialization Center (TRECC) (http://www.trecc.org/) project funded by the Office of Naval Research implemented NSGA-II with the automated parameter-setting methodology in the Data to Knowledge (D2K) software development and data mining framework created by the National Center for Supercomputing Applications (NCSA). The resulting software, called Evolutionary Multi-objective Optimizer (EMO), was further developed with funding from BP/Atlantic Richfield to Barbara Minsker, Riverglass, Inc., and Hazard Management Systems, Inc. (HMSI). HMSI created SO as a specific application of EMO to LTMO, along with DT.

Two case studies using SO were completed in July 2004. Site A had 36 sampling locations and the optimization focused on benzene, toluene, ethylbenzene, and xylene (BTEX) characterization, while Site B had 80 sampling locations and focused on benzene characterization. Each study identified roughly 23% redundancy in the sampling locations eligible for removal.

HMSI was purchased in December 2006 by Summit, which has continued to improve, develop, and refine the software and documentation in concert with BP and this ESTCP project. During this ESTCP project, several features were added or improved based on feedback from the project team. Summit is currently performing several additional case studies to further evaluate and demonstrate the effectiveness of the software; summaries of these results will be posted on the SO website.

The Summit Software is expected to provide the following benefits related to its expected applicability:

- Significant cost savings are expected by eliminating redundant data collection. Costs could be reduced for sampling labor, laboratory analysis, data management, etc. Experience suggests that eliminating redundant sampling points can save \$500-\$1,000 or more per sample in labor and analysis. Additionally, the SO will

help managers identify trade-offs among multiple monitoring objectives, including identifying when further monitoring expenditures will likely result in minimal benefits.

- New data can be given an initial assessment for significant deviations and other features of interest with a considerable reduction in labor. Currently, at most sites even a quick visual scan for a few constituents at a few key wells can require several hours for an analyst. Thorough statistical tests could take weeks of labor but can be readily performed by DT. This may enable earlier detection and correction of potential problems and faster identification of significant changes in the physical system, which results in higher certainty of attaining protectiveness. This benefit will become even greater as emerging sensor technologies produce larger volumes of data to be analyzed and more facilities move into LTM in post-closure and passive remediation scenarios.

3.2 ADVANTAGES AND LIMITATIONS OF THE TECHNOLOGY

Related software and methods for LTM design exist, as follows:

- Geostatistical packages are widely available (e.g., Surfer, Geo-EAS) but these packages create only interpolation models and do not perform data tracking or optimization.

- Based on current information, GTS (Cameron and Hunter, 2002), the 3-tiered monitoring optimization approach by Parsons (Nobel, 2003), MAROS (Aziz et al., 2003), and Cost Effective Sampling (Johnson et al., 1996) perform various spatial and temporal redundancy analyses for LTM. However, they do not perform data tracking for site-wide targets and do not use mathematical optimization. Instead, they use heuristic ("rule-of-thumb") approaches for identifying which samples to remove, which may not identify the optimal sampling plan to best meet the site-specific objectives. Moreover, most of these methods are not yet available as supported software packages (with the exception of MAROS and a limited version of GTS). These methods do not optimize based on removing groupings of samples and do not yield trade-off curves based on the results of multi-objective mathematical optimization.

- Herrera and Pinder (1998) and Rizzo et al. (2000) have used Kalman filters for LTMO and model updating. These approaches require the use of transport models for the analysis, which most DoD sites do not have. When transport models are available, Kalman filters can use available data to update the models as new data become available. Herrera and Pinder's approach uses this capability to identify the next location that should be sampled, selecting the location with the most uncertainty in the model predictions, after each sample is collected with an event. This approach assumes that data at a site are collected sequentially, with enough time between each sample to analyze the previous result and use it to determine the next sample location, which is not always the case. Additionally,

the sequential sampling approach is not a global algorithm that identifies the best set of locations to maximize the overall reduction in uncertainty. Rizzo et al. couple the Kalman filter with simulated annealing, a global optimization approach, but to our knowledge it is not available as a software package.

Some of the advantages of the Summit Software demonstrated in this ESTCP project are listed below:

- The Summit Software is the only user-friendly software available (i.e., not a research code) that performs monitoring optimization with mathematical algorithms that provide optimal or near-optimal solutions with high probability.

- A major advantage of the optimization approach utilized in the Summit Software is that it allows sampling redundancy to be evaluated on a system-wide basis, identifying optimal solutions with one, two, three, etc. locations removed, rather than on a well-by-well basis such as the redundancy analysis employed in MAROS.

- It is the only LTMO software available that enables users to select multiple site-specific monitoring objectives for the redundancy analysis, thus allowing the trade-off between the number of samples and the resulting error to be rigorously evaluated.

- Visualizations of the plume for the baseline plan with all samples versus improved plans with reduced number of samples are created within the software, which is not the case with MAROS.

- It is the only LTMO software currently available to incorporate data tracking capabilities to semi-automatically identify unexpected values in recently collected data.

On the other hand, the Summit Software does not incorporate some features that are available in other software. For example, (1) it has fewer types of interpolation models than most available geostatistical packages (although the interpolation models available in the Summit Software are standard and well accepted); (2) its data tracking routine does not include trend analysis (such as performed by MAROS) to indicate if concentration trends at individual wells are increasing, decreasing, or stable; and (3) it does not incorporate transport models as does the Kalman filter approach.

Although the Summit Software performs spatiotemporal optimization, it does so using sequential spatial interpolation and does not include intra-well temporal interpolation, which likely makes the results of the spatiotemporal optimization overly conservative (this is discussed in more detail later in the report). Another limitation is that the mass and mass flux tracking features in the software are impacted by differences in the number of sampling locations in each event. This is related to the fact that the software does not perform temporal interpolation or extrapolation to

fill in missing values in events where specific wells are not sampled. It is noted that MAROS has the same limitation in this regard.

The Summit Software is designed for sites meeting the following criteria, which represent limitations to application of the software; other approaches or software products designed for similar purposes will undoubtedly have similar limitations:

- For Model Builder and SO to provide reliable results, the site should be in an LTM situation, which may include ongoing active or passive remediation. This implies that the groundwater chemistry should be expected to change smoothly if at all in the foreseeable future, whether or not remediation activities are underway, and that no new sources or potential releases are anticipated.

- The site should have an adequate data history available that is representative of its current groundwater chemistry and status. If spatiotemporal redundancy evaluation using Model Builder and SO is to be performed, a site should generally have at least four observations obtained over a period of at least 2 years at most wells under consideration (spatial redundancy evaluation has no such requirement), and eight observations are preferred. Strictly speaking, when using SO for spatial optimization, only one data value per well for each primary COC is needed for using the software, but a more extended data history is preferable to verify the presence of the requisite LTM situation. For DT, the software requires an absolute minimum of four background observations per COC per well; more observations are desirable.

- The specific number of wells needed depends on the site complexity. In general there should be at least 20 monitoring wells in order to anticipate significant cost reductions from a spatial redundancy analysis (otherwise a significant percentage reduction in sampling locations would only yield a modest cost savings). If one is interested only in using DT, an adequate data history for each constituent of interest at each well is needed (at least four previous samples to use as background data), but there is no requirement for a minimum number of wells.

- There will generally be one, or at most a few, primary COCs with respect to which modeling and optimization for redundancy are performed. DT is not limited to the COCs used by Model Builder and SO.

- For efficient use of DT, future routine monitoring data should be made available in compatible electronic form. This may be expedited by constructing a "data bridge" mechanism (i.e., a program that reads data in one format and rewrites the data into a desired format).

- One of the assumptions for applying the Summit Software is that there are no major discontinuities in the specific aquifer being evaluated with respect to hydraulic connection. Highly fractured media would not meet that assumption. This limitation would of course be true of LTMO software in general, not just the

Summit Software. Similarly, sites with extremely large contrasts in hydraulic conductivity (e.g., preferential pathways) might impact the application of the software.

The limitations listed above pertain to the overall technology. Several additional limitations within the software functionality observed during testing are discussed later in this report.

This page left blank intentionally.

4.0 PERFORMANCE OBJECTIVES

Table 1 summarizes the performance objectives for evaluating the Summit Software provided in the Technology Demonstration Plan. To avoid repetition, a detailed discussion is provided for each performance objective in Section 7.

Table 1. Performance objectives.

Type of Performance Objective	Performance Criteria	Expected Performance (Metric)	Performance Objective Met?
Qualitative	User functionality (primary)	The Summit Software has an acceptable learning curve (e.g., 1-2 days) that will not discourage prospective users and allows users to achieve the intended objectives.	Yes
	Software reliability (primary)	The Summit Software has no significant errors or bugs remaining by the end of this project.	Yes
	Model Builder performance (primary)	Model Builder provides a model of spatial and/or temporal variation for each primary constituent of concern at each site that is adequate given the available data.	Yes (Spatial) Partially (Temporal)
	SO performance (primary)	SO provides reasonable trade-off curves, allowing site personnel and other professionals to easily identify optimal monitoring program choices.	Yes (Spatial) Partially (Spatiotemporal)
	DT performance (primary)	DT enables the easy incorporation of site-specific monitoring and remediation expectations and data objectives along with historical data.	Partially
	Regulatory acceptance (primary)	Results of this ESTCP dem/val will be persuasive to regulatory personnel.	Yes
	Comparison with MAROS (secondary)	The Summit Software will be found to be, at most, modestly more difficult to learn to use, consistent with being much more flexible in incorporating site-specific monitoring objectives. There are no preconceptions regarding the comparisons of optimization recommendations to be expected from the two software products.	Yes

Table 1. Performance objectives (continued).

Type of Performance Objective	Performance Criteria	Expected Performance (Metric)	Performance Objective Met?
Quantitative	Model Builder performance (primary)	Model Builder provides a model of spatial and/or temporal variation for each primary constituent of concern at each site that is adequate given the available data.	Yes (Spatial) Partially (Temporal)
	SO performance (primary)	Optimized programs identified by SO in fact permit cost reductions with acceptable losses of information, if appropriate, as anticipated for the large majority of DoD sites.	Yes
	DT performance (primary)	DT responds appropriately to artificially induced anomalies of interest for the particular site.	Yes
	Comparison with MAROS (secondary)	The Summit Software will be found to be, at most, modestly more difficult to learn to use, consistent with being much more flexible in incorporating site-specific monitoring objectives. If both products are able to accept the same goals and constraints, results will be similar but slightly different due to different optimization methodologies. There are no preconceptions regarding the comparisons of optimization recommendations to be expected from the two software products.	Yes

5.0 SITE DESCRIPTION

Three DoD demonstration sites were selected. Table 2 provides a summary of the demonstration sites.

Table 2. Summary of demonstration sites.

	Norfolk Naval Station Camp Allen Landfill	Former George AFB (GAFB) OU1	Former Nebraska Ordnance Plant (NOP) OU2
Agency	Navy	Air Force	Army
Location	Norfolk, VA	Victorville, CA	Mead, NE
Geographic location	East (coastal)	West (arid)	Midwest (plains)
Remediation system	Pump-and-treat (P&T) with air stripping for hydraulic containment	P&T started in 1991 and shut down since 2003	P&T with 10 extraction wells
Primary COCs	c12DCE[1], TCE[2], VC[3]	TCE	TCE and RDX[4]
Aquifers evaluated	Shallow and deep aquifers	Upper aquifer	Shallow, intermediate, and deep aquifers
Sampling frequency	Annual	Semi-annual	Varies by well
Monitoring network	~70	~50	~220

[1]cis-1,2-dichloroethene [2]trichloroethene [3]vinyl chloride [4]Hexahydro-1,3,5-trinitro-1,3,5-triazine

5.1 SITE LOCATION AND HISTORY

Locations are summarized in Table 2. Location maps are provided in the Final Report.

5.2 SITE GEOLOGY/HYDROGEOLOGY

The number of aquifers evaluated at each site is summarized in Table 2. Details regarding stratigraphy and groundwater flow patterns at each site are provided in the Final Report.

5.3 CONTAMINANT DISTRIBUTION

The primary COCs at each site and the approximate number of wells comprising the monitoring network at each site are summarized in Table 2. Plume maps are provided in the Final Report.

This page left blank intentionally.

6.0 TESTING DESIGN

6.1 CONCEPTUAL EXPERIMENTAL DESIGN

Initially, information was obtained about a number of candidate sites. Three sites were selected based on considerations that included monitoring and regulatory status, data availability, and representation across the DoD. The following general approach was applied for each of the demonstration sites selected:

- The ESTCP project team obtained preliminary information from the site team for review prior to site visit (e.g., reports describing site conceptual model and site history).

- The ESTCP project team conducted a site visit to present an overview of the project and to receive input regarding the optimization formulation from the site team.

- The ESTCP project team developed a preliminary optimization formulation, provided it to the site team for review, and then finalized the optimization formulation based on feedback from the site team.

- The site team then provided the most updated version of historical sampling data to EnviroStat in electronic format.

- EnviroStat then performed the following activities:

    - Screened historical data to determine if any obvious data quality issues were evident

    - Attempted to resolve any data quality issues with the site team

    - Reserved the last year of sampling data (which would be used later for validation of SO results as well as for evaluation of DT)

    - Provided GeoTrans with comma separated variable (CSV) files, not including the reserved data that could be used as input to the software for evaluation of the SO functionality

    - Created five alternate versions of the reserved data that incorporated artificial anomalies based on discussions with the site team about scenarios of potential interest.

- GeoTrans applied the Summit Software to evaluate data redundancy (i.e., the SO module in conjunction with Model Builder) and provided the ESTCP project team with a preliminary write-up of results.

- EnviroStat then provided GeoTrans with the six versions of the reserved (i.e., current) data for input to DT (i.e., one was the actual version, and the other five had artificial anomalies as well as tweaked values so that identification of the actual version by GeoTrans would be difficult).

- GeoTrans applied the Summit Software to evaluate the DT functionality, and provided a summary of DT results to EnviroStat.

- EnviroStat then revealed which of the six versions of the reserved data was the actual version, and GeoTrans used that dataset to perform validation of the SO results (i.e., used the more recent data to make plume maps using the baseline well locations, and plume maps based on the optimized sampling plans, to see if the maps based on the optimized sampling plans appeared reasonable).

- GeoTrans provided the ESTCP team with a write-up of results and conclusions, incorporating sections provided by EnviroStat regarding data preparation and interpretation of DT results.

- After review by the ESTCP team, the write-up was finalized and forwarded to the site team for their review and feedback, followed by a presentation of results to the site team via conference call (Camp Allen site, NOP site) or in-person meeting (GAFB site).

In addition to the analysis for each site described above, the following activities were also part of the project design:

- For one of the three sites (GAFB), additional validation of results was performed based on one year of sampling conducted subsequent to the original set of reserved data.

- For one of the three sites (Camp Allen) the MAROS software was also applied, so that functionality and results (to the extent possible) could be compared (note that GTS and MAROS will also be applied at the NOP site in a separate effort outside of this project).

- USEPA Region V arranged for their contractor to apply the software at one of their sites and provided feedback to the ESTCP project team by filling out a questionnaire and preparing a brief summary report.

Throughout the project the Summit Software was applied by a mid-level GeoTrans engineer with no previous experience with LTMO software, rather than by the software developer. In this manner, the software demonstration provided a realistic evaluation of its usability by a typical DoD contractor with no prior experience with the software. Before the first site was evaluated, the GeoTrans analyst was provided with basic training for both the Summit Software and the MAROS code. The subsequent two sites were evaluated by a different mid-level GeoTrans analyst who was not provided any training on either software.

6.2 BASELINE CHARACTERIZATION

The steps in the experimental design that might be considered baseline characterization are those associated with the developing the optimization formulation for each site and obtaining and preparing the data for each site. Those items are described in detail in the Final Report.

6.3 TREATABILITY OR LABORATORY STUDY RESULTS

These items do not apply to this ESTCP project.

6.4 FIELD TESTING

The schedule for testing of the software is summarized in Table 3.

Table 3. Schedule for testing the software.

	2007												2008											
	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D
Camp Allen Site:																								
Site visit	■																							
Formulation					■																			
Software testing							■	■	■	■														
Apply MAROS							■																	
GAFB Site:																								
Site visit			■																					
Formulation												■	■											
Software testing															■	■	■	■						
Follow-up evaluation																					■	■		
NOP Site:																								
Site visit	■																							
Formulation																	■							
Software testing																			■	■				
USEPA Testing at One Site											■	■												

The analysis for each demonstration site incorporated variations that increased the robustness of the testing. These variations included the following:

- Camp Allen Site
  - o Evaluated two aquifers (shallow and deep)
  - o Evaluated all six combinations of interpolation method and transformation type in Model Builder
  - o Evaluated all six combinations of interpolation method and transformation type for spatial redundancy evaluation, and two of the combinations (kriging-quantile and inverse distance weighting (IDW)-quantile) for spatiotemporal redundancy evaluation
  - o Evaluated three COCs simultaneously for redundancy evaluation

    - Evaluated different values for Population Size in the GA in SO for both spatial and spatiotemporal redundancy evaluation.

- GAFB Site
    - Evaluated one aquifer (upper)
    - Evaluated all six combinations of interpolation method and transformation type in Model Builder
    - Evaluated two combinations of interpolation method and transformation type for spatial redundancy evaluation (kriging-quantile and IDW-quantile) and one combination for spatiotemporal redundancy evaluation (kriging-quantile)
    - Evaluated two variations of the baseline data: Dataset A had 55 wells and Dataset B had 47 wells (eight wells that had atypical water levels and screened intervals were removed from Dataset B)
    - Evaluated different combinations of values for Population Size and Number of Generations in the GA in both spatial and spatiotemporal redundancy evaluation
    - Evaluated the Mass Metric and Mass Flux functionality for two cases: one with uneven distribution of samples per event, and one where missing data were filled in manually (outside the software) based on temporal interpolation
    - For spatiotemporal redundancy evaluation, evaluated use of the original dataset versus a modified dataset where missing data were filled in manually (outside the software) based on temporal interpolation
    - For both spatial and spatiotemporal redundancy evaluation, utilized three different values for cut-off concentration between low concentration (i.e., plume boundary) and high concentration (i.e., plume interior) areas.
- NOP site
    - Evaluated three aquifers (shallow, intermediate, and deep)
    - Evaluated all six combinations of interpolation method and transformation type in Model Builder
    - Evaluated all six combination of interpolation method and transformation type for spatial redundancy evaluation
    - Evaluated the difference between considering multiple COCs (i.e., TCE and RDX) simultaneously versus independently for the redundancy evaluation.

In addition, for the DT evaluation at each of the three sites, artificial anomalies were added to the actual reserved data by EnviroStat. These anomalies followed a number of scenarios that had been discussed with the site teams, such as

- Abnormally high concentrations in a particular area of the site due to new contaminant source and/or plume migration
- Abnormally high concentrations at individual wells for no apparent reason
- Abnormally low values for a number of samples in the same event due to a systematic lab error
- Switched samples (e.g., bottles labeled incorrectly or reported incorrectly by the lab)
- Laboratory cross-contamination
- Database errors.

In addition, it turned out that some of these anomalies were present in the actual data from two of the sites.

During the application of the software at the first two demonstration sites, several minor bugs were detected and reported to the software developer. For example, in DT an issue with plotting the y-axis data labels was identified. By the end of the demonstration product, no remaining bugs were known to exist. During this demonstration project, several software features were added or improved; those are detailed in the Final Report. These were not due to bugs; rather, they represent the evolution of software features.

The detailed results obtained from the field testing are presented in the Final Report. Some examples of the results are presented below. Figure 3 is an example trade-off curve from the GAFB spatial redundancy analysis (the details regarding how error is calculated and how the optimization problems were formulated are provided in the Final Report). Each diamond represents the optimal monitoring design for a given level of expenditure (i.e., number of wells). Such solutions are optimal (non-dominated) because no other possible solution is superior in both objectives simultaneously. The advantage of using GAs to solve multi-objective problems is that the entire set of such optimal solutions can be generated in a single optimization run. In this case the user selected two plans along the trade-off curve (plans 97 and 14) for further evaluation (comparing plume maps generated by the software for those plans to each other, and to the baseline plan with all sampling locations). The user can select any plan along the trade-off curve for further evaluation. Plans further to the left will have more wells remaining (i.e., fewer wells eliminated) and less maximum value of error at any of the removed locations.

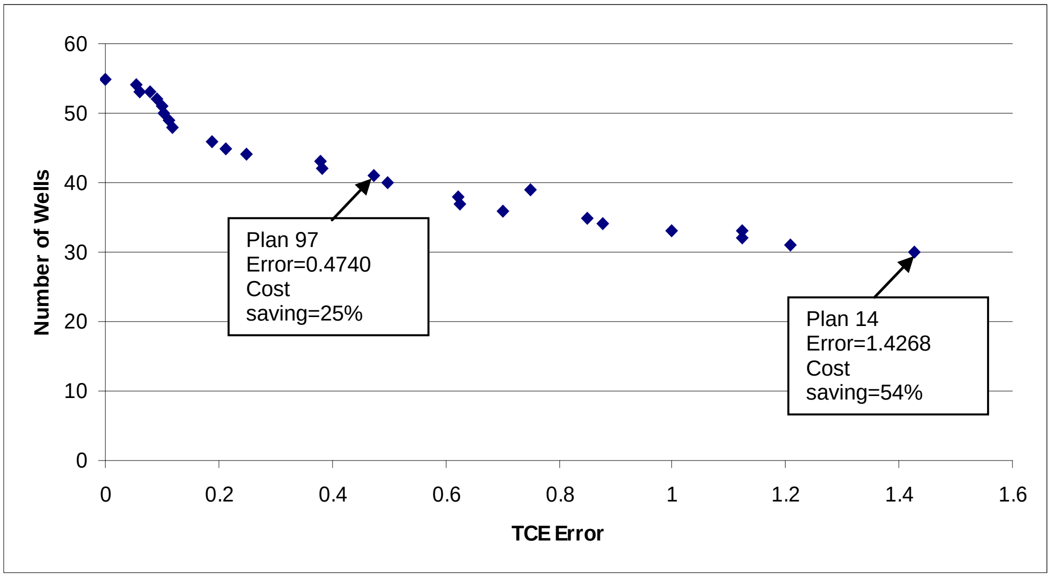
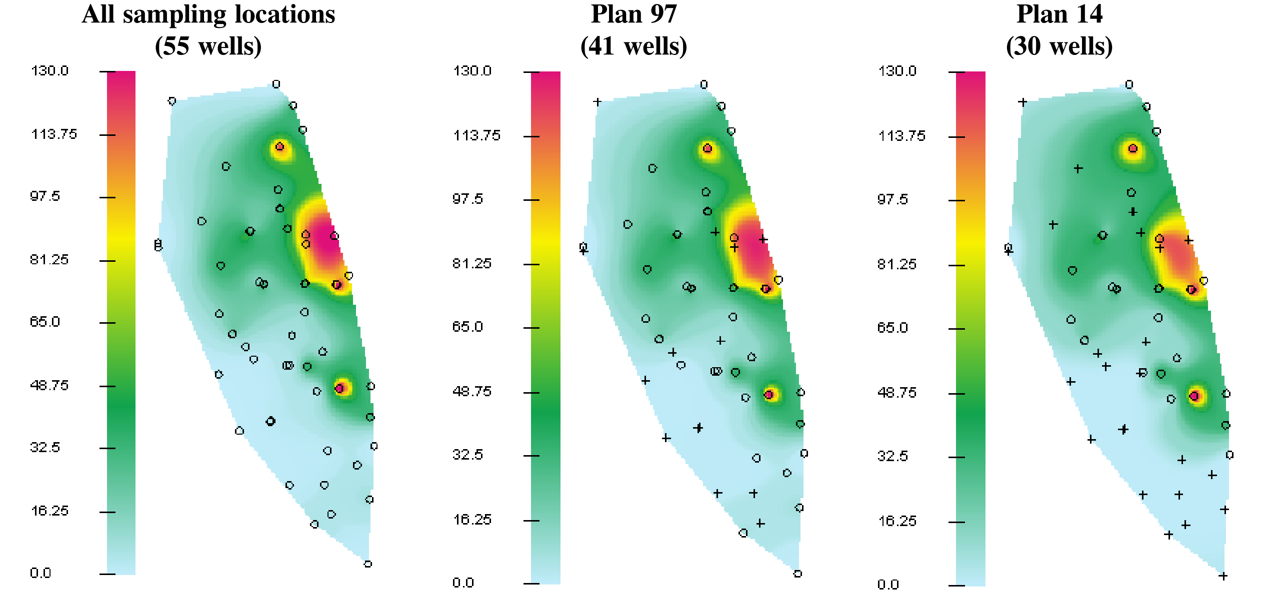


Figure 3. Example trade-off curve for spatial redundancy evaluation, GAFB site.

Figure 4 is an example of plume maps generated by the software for two plans selected from the tradeoff curve above. In this case the plume maps illustrate that the selected plans with reduced number of sampling locations provide for a generally similar graphical representation of the plume.



Note:
The symbol "+" indicates wells that are recommended to be removed by the Optimizer, while the symbol "O" denotes wells that are recommended to keep.

Figure 4. Example of plume maps for plans selected from the trade-off curve.

It is difficult to fully summarize the results for the three demonstration sites, since the software produces a trade-off curve for each optimization simulation, which provides a family of optimal solutions. In Tables 4 through 6, the savings versus the baseline data set are presented for different values of normalized error along the trade-off curve.

Table 4. Summary of redundancy evaluation results, Camp Allen site.

Optimization	# of samples in the baseline model	# of samples with the max error per COC of 0.5	# of samples with the max error per COC of 1.0	# of samples with the max error per COC of 1.5
Shallow Aquifer				
Spatial	42	17 (59.5% saving)	16 (61.9% saving)	14 (66.6% saving)
Spatiotemporal	N/A	N/A	N/A	N/A
Deep Aquifer				
Spatial	31	28 (9.7% saving)	21 (32.2% saving)	21 (32.2% saving)
Spatiotemporal	21	20.2 (3.8% saving)	17.7 (15.7% saving)	17.5 (16.7% saving)

Note: Number of samples is the number of wells (spatial) or the number of samples per year (spatiotemporal)
Aquifers: Shallow and deep (evaluated separately)
COCs evaluated: c12DCE, TCE, and VC (evaluated simultaneously, errors for each COC added together)
Model: Kriging interpolation with quantile data transformation

Table 5. Summary of redundancy evaluation results, GAFB site.

Cutoff Value	Optimization	# of samples in the baseline model	# of samples with the max error 0.5	# of samples with the max error 1.0	# of samples with the max error 1.5
Upper Aquifer with Dataset A					
25		55	40 (27.3% saving)	34 (38.2% saving)	30 (45.5% saving)
50		55	42 (23.6% saving)	37 (32.7% saving)	31 (43.6% saving)
100	Spatial	55	41 (25.5% saving)	36 (34.5% saving)	32 (41.8% saving)
25		108	102 (5.6% saving)	96.33 (10.8% saving)	93.08 (13.8% saving)
50	Spatio-	108	104 (3.7% saving)	96.33 (10.8% saving)	92.83 (14.0% saving)
100	temporal	108	102 (5.6% saving)	98.08 (9.2% saving)	94.83 (12.2% saving)
Upper Aquifer with Dataset B**					
25		47	36 (23.4% saving)	30 (36.2% saving)	28 (40.4% saving)
50		47	36 (23.4% saving)	31 (34.0% saving)	28 (40.4% saving)
100	Spatial	47	36 (23.4% saving)	32 (31.9% saving)	29 (38.3% saving)
25		92	88 (4.3% saving)	82.53 (10.3% saving)	79.08 (14.0% saving)
50	Spatio-	92	88 (4.3% saving)	82.53 (10.3% saving)	79.08 (14.0% saving)
100	temporal	92	88 (4.3% saving)	82.53 (10.3% saving)	79.33 (13.8% saving)

*Note: Number of samples is the number of wells (spatial) or the number of samples per year (spatiotemporal)
** Eight wells (MW-102, MW-104, NZ-06, NZ-10, NZ-20, NZ-30, NZ-31, and NZ-32) identified not to be representative of aquifer characteristics are excluded from Dataset B.

Cutoff value is used by the software to differentiate areas of lower concentration from areas of higher
Concentration when calculating error (errors in low concentration areas are given more weight)

Aquifers: Shallow
COCs evaluated: TCE
Model: Kriging interpolation with quantile data transformation

Table 6. Summary of redundancy evaluation results, NOP site.

Optimization	# of samples in the baseline model	# of samples with the max combined error of 0.5	# of samples with the max combined error of 1.0	# of samples with the max combined error of 1.5
Shallow Aquifer (25 wells are non-removable)				
Spatial	81	54 (33.3% saving)	50 (38.3% saving)	47 (42.0% saving)
Spatiotemporal	N/A	N/A	N/A	N/A
Intermediate Aquifer (25 wells are non-removable)				
Spatial	84	48 (42.9% saving)	43 (48.8% saving)	42 (50.0% saving)
Spatiotemporal	N/A	N/A	N/A	N/A
Deep Aquifer (22 wells are non-removable)				
Spatial	56	35 (37.5% saving)	33 (41.1% saving)	32 (42.9% saving)
Spatio-temporal	N/A	N/A	N/A	N/A

*Note: Number of samples is the number of wells (spatial) or the number of samples per year (spatiotemporal)

Aquifers: Shallow, intermediate, and deep (evaluated separately)
COCs evaluated: TCE and RDX (evaluated simultaneously, errors for each COC added together)
Model: Kriging interpolation with quantile data transformation

Tables 4 through 6 clearly indicate that, for the two sites where spatiotemporal analysis was performed, the potential savings achieved with spatial optimization was far greater than the potential savings achieved with spatiotemporal optimization (no spatiotemporal analysis was performed for the NOP site). Potential savings from spatial analysis, calculated based on number of wells eliminated, ranged from approximately 10% to approximately 67%. A value of approximately 35% appears to be representative. The ultimate savings would be even greater if some reduction in sampling frequency was implemented at some of the remaining wells (using some sort of rule to fill in values for wells not sampled in specific events to make plume maps, estimate plume mass, etc.). The potential savings from the spatiotemporal analysis, based on number of samples eliminated per year, ranges from only approximately 4% to 17% at the two sites where spatiotemporal analysis was performed. An underlying conceptual reason for this is discussed in Section 6.1.4 of the Final Report.

Figure 5 is an example of a plot provided by the DT module of the software illustrating an out-of-bounds value in the most recent sampling event that might be of concern. In this instance there are only six background data values with a statistically significant decreasing trend, so DT selects exponentially decreasing prediction bounds; these are fairly wide due to the limited amount of data available. In this case, the analyst would examine the plot and the data values and most likely decide that the TCE levels in MW-23B were by now essentially around or less than the usual reporting limit and that the only "action" appropriate would be to delete the early high values from the background data set as soon as the requisite minimum number of relevant new background data values became available.

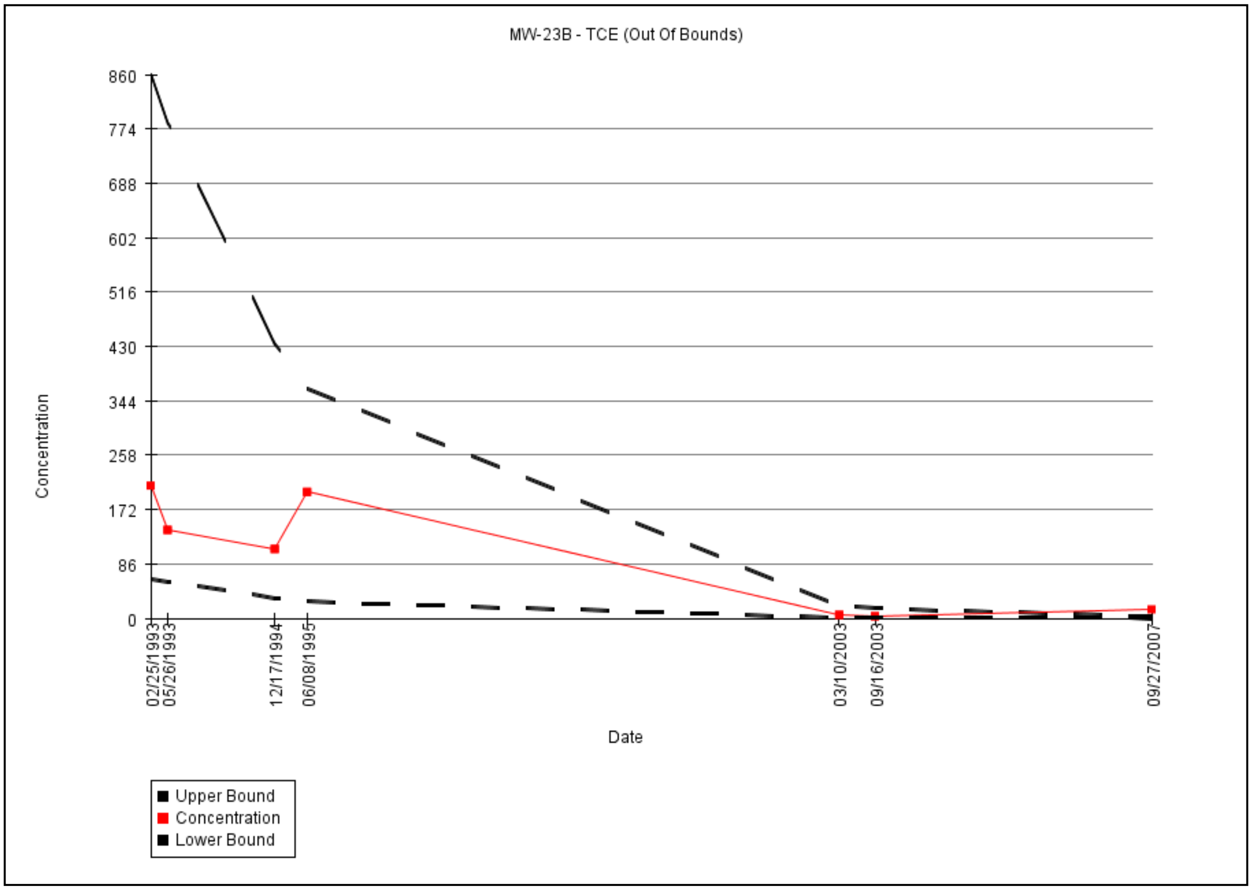


Figure 5. Example DT plot with out-of-bounds value that is a potential concern.

A summary of the amount of time it takes to apply the software is presented in Table 7. This is primarily an indication of the computation time, though the data preparation task is primarily associated with manual labor. The computation time provided for Model Builder, spatial optimization, and spatiotemporal optimization are for each problem that was simulated (e.g., each aquifer was a separate problem for the demonstration sites). Also, additional time beyond the computation time is required to interpret results.

Table 7. General summary of time required to apply the software.

Task	Time*	Comments
Data cleanup, screening, and formatting	Several days (labor)	Similar effort is needed to apply any LTMO software; effort primarily manual labor
Model builder	Minutes**	More time for kriging and higher spatial resolution
Spatial optimization	Minutes to hours**	Computation time increases with population size and generation number for GA
Spatio-temporal optimization	Hours to days**	Computation time increases with population size and generation number for GA
DT	Minutes to hours (data preparation and interpretation)	Preparation of data and plotting results requires most of the time

*For tasks where computation time is indicated, additional time is required for interpretation of results.
**Computation time per problem (e.g., per aquifer)

A more detailed summary of computation time required for the Model Builder and redundancy evaluation at the different demonstration sites is presented in Table 8.

Table 8. Summary of computation time versus number of wells.
(kriging with quantile transformation)

Site Name	Aquifer	# of samples per year	Computation Time*		
			Model Builder (kriging interpolation)	Sampling Optimizer	
				Spatial Optimization	Spatiotemporal Optimization
Camp Allen site	Shallow	42 for spatial optimization	2-4 hours**	30-60 minutes	N/A
	Deep	31 for spatial optimization 21 for spatiotemporal optimization	2-4 hours**	30-60 minutes	5-6 hours
GAFB	Upper	55 for spatial optimization 108 for spatiotemporal optimization	20 minutes	10-30 minutes	several days
NOP	Shallow	81 for spatial optimization	~5 minutes	15-50 minutes	N/A
	Intermediate	84 for spatial optimization	~5 minutes	15-50 minutes	N/A
	Deep	56 for spatial optimization	~5 minutes	15-50 minutes	N/A

N/A = not analyzed
*Pentium 4, 3.2 gigahertz (GHz)
**the programming was changed to increase speeds after the Camp Allen site was evaluated

The computation times in Table 8 are based on kriging and quantile transformation, which was determined to be the preferred model with respect to the representation of the plume. The ranges in computation time are due to some of the other parameters such as number of vertical slices (resolution) and the GA parameters such as population size. It is also noted that the programming of Model Builder was improved after the evaluation of the Camp Allen site, leading to faster computation speeds for the other two sites.

6.5 SAMPLING METHODS

No samples were collected by the ESTCP project team as part of this project. The data that were utilized were from sampling results previously obtained by the demonstration sites under their site-specific sampling plans.

6.6 SAMPLING RESULTS

Again, no samples were collected by the ESTCP project team as part of this project. The data that were utilized were from sampling results previously obtained by the demonstration sites under their site-specific sampling plans.

7.0 PERFORMANCE ASSESSMENT

7.1 QUALITATIVE PERFORMANCE OBJECTIVES

7.1.1 User Functionality

The expected performance metric is that the Summit Software has an acceptable learning curve (e.g., 1-2 days) that will not discourage prospective users and will allow users to achieve the intended objectives. For the first site, a mid-level analyst was provided a one-day training session by the software developer prior to using the software. For the second and third demonstration sites, a different mid-level analyst was used, and that person relied only on the software documentation plus phone support by the software developer. Based on the application of the software at all three demonstration sites, this performance objective was met.

7.1.2 Software Reliability

The expected performance metric is that the Summit Software has no significant errors or bugs remaining by the end of this project. Any bugs that were identified were reported to the software developer, who then fixed the problems such that the performance objective was met.

7.1.3 Model Builder Performance

The expected performance metric is that Model Builder provides a model of spatial and/or temporal variation for each primary COC at each site that is adequate given the available data. Model Builder allows for two different interpolation techniques and three different data transformation techniques, for a total of six combinations. The reasonableness of all six combinations was evaluated at each of the three sites. In each case, the analyst reviewed the plume map generated by the software for each combination of interpolation and transformation, and qualitatively assessed the reasonableness of the plume map.

The qualitative performance objective for Model Builder was met with respect to modeling of spatial variation. With respect to spatial variation, the analyst established that the combination of kriging with quantile transformation qualitatively provided the most reasonable representation of the data, which is consistent with the recommendations in the software manual. In general, using no data transformation resulted in unacceptably poor representation of the data.

The qualitative performance objective for Model Builder was only partially met with respect to modeling of temporal variation. Model Builder addresses temporal changes in concentration in support of several aspects of software functionality, such as:

- Spatiotemporal optimization for reducing sampling redundancy (in both space and time)

- Calculation of relative mass or mass flux in different sampling events, based on interpolation of sampling results in each event.

It became apparent during implementation of the software at the demonstration sites that, for each of these software functions listed above, Model Builder performs a series of spatial interpolations over time but does not perform any interpolation with respect to time. This is easily explained with a conceptual example, illustrated in Figure 6.

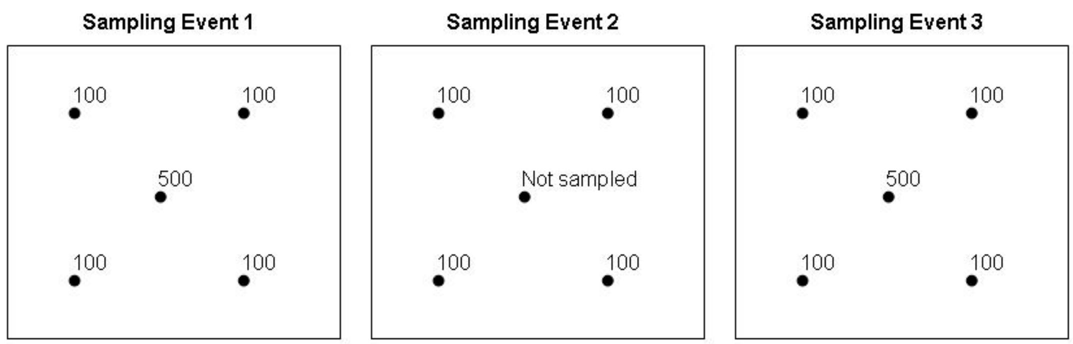


Figure 6. Conceptual example to illustrate temporal interpolation issue.

In Figure 6, the only difference between the three sampling events is that the middle point is not sampled in Sampling Event 2. The software only performs spatial interpolation within each event. Therefore, in Sampling Event 2, the software will interpolate values inside the four actual samples based only on those four values. Since those values are each 100, presumably the software will interpolate a value of 100 at each location inside the four actual values. However, it seems unlikely that an environmental scientist provided with the data illustrated in Figure 6 would consider that the most reasonable approach. Rather, a temporal interpolation at the location not sampled in Sampling Event 2 based on the other two events would appear to be more appropriate, or perhaps some combination of temporal and spatial interpolation. More specifically, an environmental scientist would likely conclude that that the value at the location not sampled in Sampling Event 2 is probably close to 500, and then perform subsequent spatial interpolation in the remaining area accordingly.

Because of this issue, the GeoTrans analyst concluded that the qualitative performance objective regarding Model Builder is only partially met with respect to temporal variation. Temporal variation is modeled adequately only if the sampling locations are consistent between sampling events, or if the user fills in missing values in specific events manually prior to use of the software, based on considerations including temporal interpolation. Otherwise, the representation of temporal variability may not be adequately modeled within Model Builder.

7.1.4 Sampling Optimizer Performance

The expected performance metric is that SO provides reasonable trade-off curves allowing site personnel and other professionals to easily identify optimal monitoring program choices. Based on the application of the software at all three demonstration sites, the qualitative performance objective for SO was met for spatial optimization but was only partially met for spatiotemporal optimization. In the case of spatiotemporal optimization, the GeoTrans analyst noted that the results seemed overly conservative because the software tended not to eliminate wells entirely from the recommended sampling plans. This was potentially not reasonable given the

corresponding spatial optimization results, which suggested that many wells could be eliminated with acceptable levels of error. Details regarding specific results for demonstration sites that pertain to this issue were presented in the Final Report. However, further explanation of an underlying issue associated with spatiotemporal optimization is provided below.

Spatiotemporal results are inherently more conservative than the spatial results alone, because the optimization problem has additional constraints. However, the fact that the software does not perform temporal interpolation of data provides an additional component of conservatism.

The reason that the spatiotemporal optimization can be overly biased towards not eliminating wells is illustrated using a conceptual example presented in Figure 7. This figure illustrates data values from three sampling events.

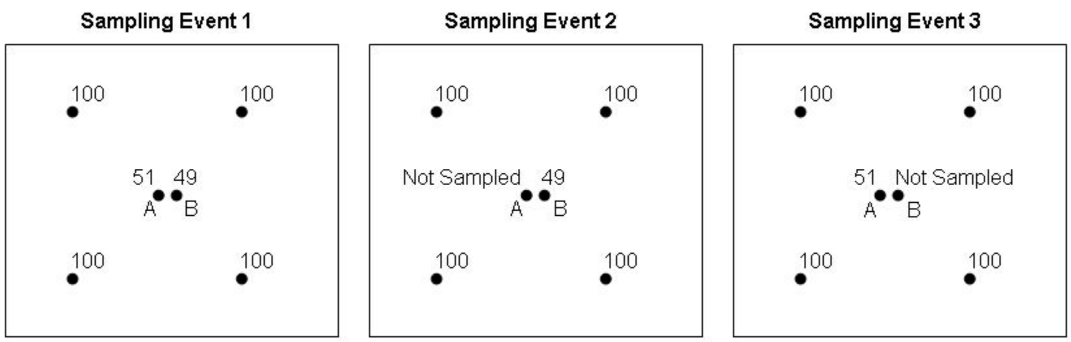


Figure 7. Conceptual example to illustrate spatiotemporal optimization issue.

On Figure 7 note the locations labeled “A” and “B” in the middle of each event. In Sampling Event 1, both those locations are sampled. In Sampling Event 2, location “A” is not sampled, and in Sampling Event 3, location B is not sampled. It appears based on inspection that locations A and B appear to be redundant. It is likely that the missing value at A in Sampling Event 2 is close to 51, and likely that the missing value at B in Sampling Event 3 is close to 49. Furthermore, it would not cause significant interpolation error if either A or B was permanently eliminated, as long as a value at the other well was available.

However, the software would likely not reach the same conclusion. Spatiotemporal optimization, as implemented in the software, provides a trade-off curve of sampling cost versus sampling error. In a spatiotemporal optimization, the error calculated by the software for each potential sampling plan is the maximum error at any sampling location in any sampling event, calculated at locations/times where a sample is removed, based on spatial interpolation using the remaining samples in that event. As discussed with respect to Model Builder, no temporal interpolation is performed within the software to address the missing values as part of these calculations. Thus, in the example in Figure 7, the software would generally not eliminate location A because it would lead to a large error in Sampling Event 3 at location A (because no value would be assumed at location B during that event). Similarly, the software would not eliminate location “B” because it would lead to a large error in Sampling Event 2 at location B (because no value would be assumed at location A during that event). Thus, the ultimate result

from the software would likely include continued sampling at both wells at some frequency, rather than eliminating one of the two wells.

This underlying issue will be significant only when well locations are not consistent for each sampling event. However, it is very often the case that at least some sampling events will have different sampling locations than others (due to different sampling frequencies, new wells, abandoned wells, and wells that cannot be sampled in specific events due to logistics). Because the software does not provide for temporal interpolation, those missing values are not adequately represented within the spatiotemporal optimization process. The assignment of these missing values can be assigned external to the software, but that is potentially a labor-intensive process. Moreover, assigning values to missing data using any form of interpolation involves making subjective judgments about the temporal smoothness of the data; as a result, a temporal optimization will be evaluating some combination of the actual temporal redundancy and the temporal redundancy of the subjective judgments, rather than evaluating only the actual temporal redundancy. It seems likely that this issue would be problematic to any approach to temporal optimization that one might attempt to use with datasets containing numerous missing values.

Therefore, in summary, the software does produce reasonable trade-off curves for spatial optimization, and in some cases where sampling locations are not consistently sampled, the trade-off curves for spatiotemporal optimization may be overly conservative with respect to elimination of wells.

7.1.5 Data Tracker Performance

The expected performance metric is that DT enables the easy incorporation of site-specific monitoring and remediation expectations and data objectives along with historical data. The qualitative performance objective for DT was partially met. DT was found to be easy to use. DT does allow tracking of recently collected concentration data against statistically computed bounds that are calculated from the historical (i.e., background) data. The software indicates which of the recently collected data values are out-of-bounds relative to expectations that are calculated from the background values. However the software does not appear to track any quantity with respect to remediation expectations or site-specific objectives, as stated in the performance objectives. With respect to concentrations, the software does not indicate which values are above or below remediation goals. With respect to other parameters that can be tracked by the software such as mass and mass flux, the software will provide a table of historical versus current values, but it does not compare these to any site-specific remediation goals for those parameters. Therefore, the software does effectively track recently collected data versus historical data, and does allow quantities such as relative plume mass and mass flux across a site boundary to be tracked, but does not specifically address "site-specific remediation expectations" within the software. In fairness, one should note that tracking progress toward site-specific remediation goals did not arise in discussions of objectives with site personnel at any of the three demonstration sites.

7.1.6 Regulatory Acceptance

The expected performance metric is that the results of this ESTCP dem/val will be persuasive to regulatory personnel. Site personnel for the demonstration sites indicated that the types or trade-

off curves produced by the software for evaluating redundancy (based on mathematical optimization), in conjunction with the comparison of plume visualizations with and without redundant data that are produced by the software, would be convincing.

7.1.7 Comparison with MAROS (secondary)

The expected performance metric is that the Summit Software will be found to be, at most, modestly more difficult to learn to use compared to MAROS. This was evaluated in this project by applying both software products at the first site, conducted by the same analyst, with equivalent training in both software products. Based on this effort, the performance objective was met. The analysts reported that both software products were equally easy to learn and implement. In the case of applying MAROS, there was some initial confusion about how to utilize some of the input parameters (current plume length, distance from source, etc.) for the specific problem being solved, but those were easily addressed by corresponding with the software developer. Similarly, there were some minor questions about what values to assign for some of the Summit Software parameters, and those were easily addressed by corresponding with the software developer (some of those issues were subsequently addressed by improvements made to the software and added to the Users Guide during the remainder of the project). More detail regarding the comparison between the Summit Software and MAROS is presented in Appendix F of the Final Report.

7.2 QUANTITATIVE PERFORMANCE OBJECTIVES

7.2.1 Model Builder Performance

The expected performance metric is that Model Builder provides a model of spatial and/or temporal variation for each primary constituent of concern at each site that is adequate given the available data. This was evaluated in this project by testing the various Model Builder options at the three demonstration sites and making visual comparisons. It is not clear that there is a more quantitative manner to evaluate this objective. Thus, the conclusions are the same as stated for the qualitative performance objective for Model Builder (see Section 7.1).

7.2.2 Sampling Optimizer Performance

The expected performance metric is that the optimized programs identified by SO in fact permit cost reductions with acceptable losses of information, if appropriate, as anticipated for the large majority of DoD sites. This performance objective was met. At each of the three demonstration sites, sampling plans were selected from the trade-off curves (sampling cost versus error) with substantially lower cost that the analyst felt had acceptable error. Furthermore, at each of the three sites validation was performed on a subsequent data set to determine if the errors that might result from the reduced sampling in the recommended plans were acceptable, and in each case it was confirmed that a plan recommended by the software with significantly reduced sampling resulted in acceptable error when applied to the reserved dataset. The validation exercise provided confidence in the results provided by SO.

7.2.3 Data Tracker Performance

The expected performance metric is that DT responds appropriately to artificially induced anomalies of interest for the particular site. This was evaluated in this project by having EnviroStat create artificial anomalies in consultation with site and other personnel at each of the three demonstration sites. Those data were provided to the software analyst from GeoTrans who used DT in a single-blind evaluation (i.e., the GeoTrans analyst was provided six variations of the actual current data with no external clues as to which was correct and which had artificial anomalies). EnviroStat subsequently evaluated the extent to which the GeoTrans analyst was able to detect these artificial anomalies with the software and provided a summary write-up for each site. Based on those efforts, this performance objective was met. The vast majority of artificial anomalies were detected (details provided in the Final Report).

7.2.4 Comparison with MAROS

The expected performance metric is that: 1) the Summit Software will be found to be at most modestly more difficult to learn to use, consistent with being much more flexible in incorporating site-specific monitoring objectives; and 2) if both products are able to accept the same goals and constraints, results will be similar but slightly different due to different optimization methodologies. There were no preconceptions regarding the comparisons of optimization recommendations to be expected from the two software products.

With regard to ease of use, GeoTrans personnel received training for both the Summit and MAROS tools by the software developers. The Summit training was a little more than a half-day, and the MAROS training was approximately a half-day hands-on training that was part of a two-day LTMO conference. Both software products were similarly easy to install and could be learned and used by people with similar training and qualifications. Both Users Guides are comprehensive and clearly presented. Similarly, to get historical site data into the required format for input to the software was no more significant for one software versus the other software; the user simply needs to follow the instructions regarding the input structure and requirements for that software product. These modifications took on the order of minutes to several hours for each software product for someone experienced with performing such operations in Microsoft Excel or Microsoft Access.

With regard to results, comparisons of specific recommendations provided by each software package are difficult, as explained in detail in Appendix F of the Final Report. Key comparison observations that can be made include the following:

- The primary advantage of the Summit Software is that the redundancy evaluation is based on mathematical optimization, which allows sampling redundancy to be evaluated on a system-wide basis (e.g., best solution if one location is removed, if two locations are removed, if three locations are removed, etc.). A key benefit of this approach is that it allows the trade-off between the number of samples and the accuracy of the resulting plume interpolation to be assessed. This is a significant improvement over the approach for evaluating data redundancy utilized in MAROS, which is not based on mathematical optimization. In MAROS, individual well locations are evaluated for redundancy based on impacts of

removing that well alone; consequently, the impact of removing groups of wells cannot be assessed and the aforementioned trade-off cannot be evaluated.

- The Summit Software approach to data redundancy evaluation provides plume visualizations for the baseline plan (i.e., all samples) versus improved plans (i.e., reduced numbers of samples) within the software. These comparative visualizations are quite effective for communication with stakeholders and regulators. However, these maps can only be exported as image files; thus, it is difficult to directly import these image files into other software packages such as Surfer, ArcGIS, and AutoCAD. MAROS does not include such plume visualizations.

- The Summit Software has a DT module that indicates if new data are “in-bounds” or “out-of-bounds” relative to expectations, based on previous data at that well. This functionality is useful but is not present in MAROS. MAROS indicates if the concentration trend at a well is increasing, decreasing, or stable. This functionality is also useful but is not present in the Summit Software.

- Both software products suffer from some similar limitations. With respect to areas of uncertainty, neither software package provides specific recommendations (i.e., number of new wells and/or locations of new wells) to reduce the uncertainty. With respect to mass calculations, neither software performs temporal interpolation or extrapolation to fill in missing values in events where specific wells are not sampled. As a result, mass or mass flux results will have higher variability and uncertainty for events with fewer samples, different spatial distribution of samples, and/or for events where key wells (e.g., wells with high concentrations) are not sampled.

Additional comparison with MAROS is planned using data from the NOP site, but that comparison will be performed outside the scope of this ESTCP project.

This page left blank intentionally.

8.0 COST ASSESSMENT

8.1 COST MODEL

The software is free for use at DoD sites. Furthermore, since the software runs on standard desktop computers, no capital purchases are required. Therefore, the cost of implementation is the estimated cost of applying the software at a typical site, and perhaps some minor training costs for initial use. For the demonstration project, approximately $60,000 per site was allocated for testing the software. However, this is far in excess of what would be required for a typical site. This is because many potential variations were addressed during this project to allow for robust testing of the software. In Table 9, estimates are provided for applying the software at a typical DoD site assuming that the redundancy evaluation will be performed spatially rather than spatiotemporally.

Table 9. Estimated costs to apply the software at a typical DoD site.

Cost Element	Estimated Level of Effort	Estimated Cost
Start-Up		
Software cost	Free	$ 0
Software download	1 hr @ $100/hr	$ 100
Training/learning	16 hrs @$100/hr	$1,600
	Subtotal	**--------->$1,700**
Redundancy Evaluation (Periodic)		
Per Site:		
Formulation	Lump sum	$5,000
Data prep	24 hrs $100/hr	$2,400
Import data into software	2 hrs @$100/hr	$ 200
	subtotal	**--------->$7,600**
*Per Plume Evaluated**:		
Model Builder	2 hrs @ 100/hr	$ 200
Optimization **	24 hrs @ 100/hr	$2,400
Interpret results and write up	20 hrs @ 100/hr	$2,000
	subtotal	**--------->$4,600 (per plume)**
Data Tracker		
First Time:		
Develop initial background data file	(Part of Data Prep listed above)	$ 0
Each Year:		
Evaluate need to update background	16 hrs@100/hr	$1,600
	subtotal	**--------->$1,600 (per year)**
Each Event:		
Create CSV file for new data	2 hrs @100/hr	$ 200
Import data and run DT	1 hrs @100/hr	$ 100
Export charts, print charts, interpret	5 hrs @100/hr	$ 500
	subtotal	**--------->$800 (per event)**

*Each plume may consist of multiple primary COCs, but each aquifer or aquifer horizon where the plume is represented with a different map would be treated as separate plume.
**Assumes several variations will be attempted such as changing the Model Builder algorithm or the list of excluded wells.

8.2 COST DRIVERS

The cost estimates provided in Table 9 are rough estimates based on the testing performed as part of this demonstration project. Some cost drivers that would potentially impact the cost of applying the software are provided below:

- Formulation task will depend on number of people and need for additional meetings.
- Data preparation cost will depend on the quality of the site data.
- Redundancy evaluation costs depend on the number of plumes.
- Spatiotemporal optimization requires much more computation time than spatial optimization.

The items above are discussed in detail in the Final Report. Also, the computation time is somewhat impacted by the number of wells and the model type (kriging versus inverse distance weighting), but these variations should not significantly impact the costs estimated in Table 9. It is also noted that the labor cost estimates in Table 9 are approximations that may differ from site to site.

8.3 COST ANALYSIS

A cost-benefit analysis for applying this LTMO software must account for the costs of applying the technology and the cost savings likely to be realized. The estimated costs of applying the technology were presented in Table 9. The costs savings will result from reduced labor and analysis associated with the elimination of some sampling.

The actual costs and savings are subject to many site-specific factors such as the number of aquifers, the number of wells, the cost of sampling, the cost of laboratory analysis, and many other factors. Since these factors vary from site to site, examples are provided below to illustrate how the costs and savings can be evaluated.

For the first scenario, the following assumptions are made:

- Evaluate a 10-year monitoring horizon.
- The LTMO costs are based on the values estimated in Table 9.
- There is only one aquifer, and the redundancy evaluation is performed on one comprehensive plume.
- The redundancy evaluation is performed once at the beginning of the 10-year period, and again in year 6 of the 10-year period. To support the second periodic evaluation, the full set of monitoring wells is sampled for one of the two events in year 5 (i.e., year 5 has only half the savings associated with reduced amount of wells sampled). For simplicity, we assume the same level of sampling reduction after this second round of optimization versus the original baseline number of wells.

- The wells are sampled twice per year.
- There are 60 total samples per sampling event in the current monitoring plan (i.e., 120 samples per year).
- The cost of collecting a sample, plus the laboratory cost for analysis, is $800.
- The spatial LTMO analysis eliminates 35% of the sampling locations (representative results for spatial optimization for the three demonstration sites, as described in Section 6 and presented in Tables 4 through 6).
- Future costs are discounted to present day dollars using a 10-year discount rate of 2.6% as per the Office of Management and Budget (OMB) (www.whitehouse.gov/omb/circulars/a094/a94_appx-c.html).

Other minor savings might occur because fewer duplicate samples and quality assurance/quality control (QA/QC) samples (e.g., trip blanks and field blanks) may be required, but those details have not been included. Also, the additional costs of evaluating the current data for unexpected values without using DT are hard to quantify and are not included.

The cost benefit analysis for this scenario is summarized in Table 10.

Table 10. Cost-benefit analysis, Scenario 1.

Scenario 1

	Assumptions
1	# of plumes
5	Frequency of periodic evaluation (in years)
2	Sampling frequency (per year)
60	Total # of samples per event (all plumes combined)
$800	Cost per sample (labor plus analysis)
35%	Savings for redundancy analysis
2.60%	Discount rate (OMB 10 year value)
10	Monitoring time frame (in years)

	Year										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Total**
LTMO Costs:											
Start-up	$1,700										$1,700
Periodic Evaluation	$12,200					$12,200					$24,400
Data Tracker	$1,600	$3,200	$3,200	$3,200	$3,200	$3,200	$3,200	$3,200	$3,200	$3,200	$30,400
Subtotal (non-discounted)	$15,500	$3,200	$3,200	$3,200	$3,200	$15,400	$3,200	$3,200	$3,200	$3,200	$56,500
Subtotal (discounted)	$15,500	$3,119	$3,040	$2,963	$2,888	$13,545	$2,743	$2,674	$2,606	$2,540	**$51,617**
LTM Savings:											
Optimized Sampling Plan	$33,600	$33,600	$33,600	$33,600	$16,600	$33,600	$33,600	$33,600	$33,600	$33,600	$319,200
Subtotal (discounted)	$33,600	$32,749	$31,919,	$31,110	$15,161	$29,553	$28,804	$28,074	$27,363	$26,669	**$285,001**
										Net benefit (discounted)	**$233,384**

Notes:
Costs are based on estimates in Table 9.
Costs for start-up assumed to be $1,700 for training and learning
Costs for periodic evaluation assumed to be $7,600 per site plus $4,600 per plume evaluated.
Costs for DT analysis assumed to be $800 per event plus $1,600 per year to assess need to update background data.
Potential savings from reduction in number of trip blanks, field blanks, duplicate samples, etc. not included.
Less savings assumed in year 5 – assume sampling at all wells in one of the two events to use for periodic redundancy evaluation

In Table 10 the LTM savings for the optimized sampling plan is calculated as follows:

- For most years, the savings is $33,600. This is calculated as:

  60 samples/event * 2 events/year * $800/sample * 35% reduction in samples

- For year 5, it is assumed that the full set of wells is sampled in one of the two events to provide data for a periodic redundancy evaluation, so the savings in that year is only half the amount as in the other years.

In Table 10, the net present value of the LTMO costs is approximately $50,000, and the net present value of the LTM savings is approximately $285,000. Thus, the net savings is more than $230,000 over 10 years. This relatively modest savings is because this scenario includes only one aquifer with a total of 60 wells and a relatively low cost of $800 per sample for sampling labor, analysis, and validation.

A second scenario is the same as the first scenario except for the following:

- There are three aquifers instead of one, each with 60 wells, which has the following ramifications:
  - The redundancy analysis includes three plumes rather than one plume.
  - The total number of samples per event in the current monitoring plan is 180 instead of 60 (i.e., three aquifers rather than one aquifer).
- The cost of collecting a sample, plus the laboratory cost for analysis, is $1400 rather than $800 (perhaps more parameter types to be analyzed and/or more difficult sampling conditions).

In this case the cost of the LTMO evaluations are slightly higher (since there are three aquifers rather than one aquifer). However, the estimated LTM savings for an optimized plan are much greater, since the baseline plan has many more wells and the cost per sample is higher.

The cost-benefit analysis for this scenario is summarized in Table 11.

Table 11. Cost-benefit analysis, Scenario 2

Scenario 2

	Assumptions
3	# of plumes
5	Frequency of periodic evaluation (in years)
2	Sampling frequency (per year)
180	Total # of samples per event (all plumes combined)
$1,400	Cost per sample (labor plus analysis)
35%	Savings for redundancy analysis
2.60%	Discount rate (OMB 10 year value)
10	Monitoring time frame (in years)

	Year										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Total**
LTMO Costs:											
Start-up	$1,700										$1,700
Periodic Evaluation	$21,400					$21,400					$42,800
Data Tracker	$1,600	$3,200	$3,200	$3,200	$3,200	$3,200	$3,200	$3,200	$3,200	$3,200	$30,400
Subtotal (non-discounted)	$24,700	$3,200	$3,200	$3,200	$3,200	$15,400	$3,200	$3,200	$3,200	$3,200	$74,900
Subtotal (discounted)	$24,700	$3,119	$3,040	$2,963	$2,888	$21,637	$2,743	$2,674	$2,606	$2,540	**$68,909**
LTM Savings:											
Optimized Sampling Plan	$176,400	$176,400	$176,400	$176,400	$88,200	$176,400	$176,400	$176,400	$176,400	$176,400	$1,675,800
Subtotal (discounted)	$176,400	$171,930	$167,573	$163,326	$79,594	$155,154	$151,222	$147,390	$143,655	$140,014	**$1,496,257**
										Net benefit (discounted)	**$1,427,348**

Notes:
Costs are based on estimates in Table 9.
Costs for start-up assumed to be $1,700 for training and learning
Costs for periodic evaluation assumed to be $7,600 per site plus $4,600 per plume evaluated.
Costs for DT analysis assumed to be $800 per event plus $1,600 per year to assess need to update background data.
Potential savings from reduction in number of trip blanks, field blanks, duplicate samples, etc. not included.
Less savings assumed in year 5 – assume sampling at all wells in one of the two events to use for periodic redundancy evaluation

In Table 11 the LTM savings for the optimized sampling plan is calculated as follows:

- For most years, the savings is $176,400. This is calculated as:

  180 samples/event * 2 events/year * $1400/sample * 35% reduction in samples

- For year 5, it is assumed that the full set of wells is sampled in one of the two events, to provide data to for a periodic redundancy evaluation, so the savings in that year is only half the amount as in the other years.

As summarized in Table 11, the net present value of the LTMO costs for this scenario is approximately $70,000 and the net present value of the LTM savings is approximately $1,500,000. Thus, the net savings is more than $1.4 million over 10 years. This is a very substantial net benefit.

As stated earlier, many of the parameters used in these scenarios will vary from site to site. The cost analysis approach summarized in Table 10 and Table 11 can be applied to any such set of parameters. This simple spreadsheet approach can be used to screen sites for potential benefits that might be realized from applying the LTMO software. For instance, for sites with few monitoring locations and infrequent sampling, the potential savings will be limited. However,

the cost-benefit examples provided above clearly indicate that net savings of millions of dollars are possible across the universe of DoD sites.

9.0 IMPLEMENTATION ISSUES

9.1 SOFTWARE AVAILABILITY AND DOCUMENTATION

The anticipated end users for SO and DT include government personnel and support contractors managing groundwater monitoring programs. A copy of the software executable and Users Guide is available on the Summit SO website (http://www.samplingoptimizer.com/) for free and immediate download by government employees and educational users (those accessing with ".gov," ".mil," and ".edu" extensions). Input data files from this project that can be used as sample data have also been included on the website. This website will also be linked to the ESTCP and Federal Remediation Technology Roundtable websites.

The Summit website will provide a form for contractors to government sites to fill out to obtain a license file and download link for the software. Contractors will be required to provide evidence that the software will be used at a government site (e.g., a government work order or letter from government personnel), and the license will limit the software to be able to work only with data from that site. Also, contractors will need to renew the software license annually for continuing use of the software. The free software license does not include technical support or training, which can be purchased separately (further information is available on the Summit website). Other private sector users will be able to purchase a commercial license to the software as needed. Note that this procedure is similar to those employed for other software packages such as RACER and GMS.

9.2 EASE OF USE

The software was found to be easy to use, based on the application of the software by a mid-level analyst at GeoTrans with no LTMO experience. This was true for a mid-level analyst who received training on the use of the software (for one of the three demonstration sites), as well as for a mid-level analyst who did not receive training on the software (for two of the three demonstration sites).

In addition, the USEPA group that applied the SO functions of the software (including Model Builder) outside of our project reported that: "The user interface was very easy to use…User's manual was an excellent reference for set-up and execution, and it contained clear directions for navigating dialog boxes, setting parameters, formatting input files, etc…It took only few hours to get comfortable using the software (import/export, model set up, running the program). The users manual was very helpful in this aspect. It took a few days to fully understand the method, the effects of changes in parameter values, and the results."

9.3 KEY LIMITATIONS OF THE CURRENT SOFTWARE

The software has some limitation that will impact the use of the software by end users. Key limitations (which have already been discussed in previous portions of this report) are indicated below.

- The software interpolates spatially but does not perform interpolations in time. This impacts the tracking of mass and/or mass flux when the distribution of

sampling is not consistent from event to event. It also impacts the performance of spatiotemporal redundancy analysis, resulting in more conservative results than spatial redundancy analysis when the sampling locations are not consistent from event to event. The software would be improved if there was a feature to optionally fill in missing values via temporal interpolation.

- In DT, the plots of concentration versus time do not use different symbols to differentiate between the "background data" and the "current data." The software would be improved if different symbols were used.

- In DT, the software does not allow specific historical values to be imported and plotted on graphs but not used for calculation of the prediction limits. If some historical values are considered potentially anomalous, those values have to either be included as background data (such that prediction limits are impacted) or completely ignored. The software would be improved if such values could be imported with a flag so that they can be included on concentration versus time plots (with a different symbol) but not used to calculate the prediction limits.

- The DT portion of the software does a very good job of identifying unexpected values but does not indicate whether the concentration trend for a specific COC at a specific well is increasing, stable, or decreasing. The software would be improved if that functionality was added.

- The software does not include data consolidation or recognition of flags (e.g., for non-detect values). This requires the user to consolidate the data into sampling events during preparation of the SO input files and to assign "graphing values" for non-detects during preparation of the input files for SO and DT. The software could be improved if this type of functionality was included within the software.

- Plume visualization for both Model Builder and Optimizer also allows users to change the zoom scale and color scale. The color scale is a linear scale allowing users to define the minimum and maximum concentrations for each COC. Then the software can plot the plume maps in color based on the minimum and maximum concentrations defined. However, it does not provide an option for a logarithmic scale; thus, for sites with a very big range in concentration, it cannot plot both high-end concentrations and low-end concentrations with sufficient detail (though multiple plots with different ranges could be made independently).

The first bullet listed above has the most profound implications for future use of the software. For instance, since there will generally be an uneven distribution of sampling in different events, the use of spatiotemporal redundancy evaluation may not be advisable in most cases because the results will be more conservative than those obtained using spatial redundancy evaluation (with respect to elimination of wells). An additional consideration is that spatiotemporal redundancy evaluation requires far greater computation time than spatial redundancy evaluation (for the demonstration sites, it required days for spatiotemporal simulations versus hours for spatial

simulations). Thus, a prudent approach to applying the software for reduction of redundancy might be as follows:

- Perform spatial optimization rather than spatiotemporal optimization
- Determine if eliminated well locations in one or more of the recommended plans are reasonable and acceptable
- Qualitatively specify a sampling frequency for remaining locations, based on where changes in concentration are expected and/or are of greatest concern
- Develop rules for estimating the values at locations not sampled in a specific event for developing plume maps and/or for performing mass calculations (e.g., latest value, moving average of latest values, etc.).

This approach allows the user to utilize the most powerful and beneficial aspect of the software, which is the application of mathematical optimization in conjunction with multiple objectives to develop a trade-off curve for evaluating spatial redundancy.

9.4 REGULATORY ISSUES

Regulatory approval regarding the implementation of LTMO results provided by the software primarily pertains to the results of redundancy evaluation (i.e., the SO results). Interaction with regulators regarding implementation of results at the three demonstration sites was not a specific part of this ESTCP project. Site personnel for the demonstration sites indicated that the types or trade-off curves produced by the software for evaluating redundancy (based on mathematical optimization), in conjunction with the comparison of plume visualizations with and without redundant data that are produced by the software, would be expected to be convincing.

Obtaining regulatory acceptance of the software will require two major steps: 1) increasing awareness of LTMO in general, and awareness of this software in particular, within the regulatory community and 2) making site-specific requests to regulators for modifying an LTM program based on results of the software. The project team has offered to assist each of the demonstration sites with regulatory issues associated with LTMO, but no such assistance has been requested to date. For example, the site team at the GAFB site indicated they would like to perform further analysis on their own using the software before presenting results to regulators in the form of a revised LTM plan. Also, given the long schedule of our project and the fact that the most recent data at each site were reserved for validation in our project, the sites would be advised to repeat the analyses using up-to-date data before incorporating the results into an LTM program revision proposal. Obtaining "regulatory acceptance" of the software will ultimately require that LTM modifications based on software recommendations be brought before site-specific regulators.

This page left blank intentionally.

10.0 REFERENCES

Aziz, J.J., M. Ling, H.S. Rifai, C.J. Newell, and J.R. Gonzales. 2003. MAROS: A Decision Support System for Optimizing Monitoring Plans. Ground Water vol.41, no. 3, pg. 355-367.

Cameron, K., and P. Hunter. 2002. Using Spatial Models and Kriging Techniques to Optimize Long-Term Groundwater Monitoring Networks: A Case Study. Environmetrics, 13, 629-656.

Deb, K., A. Pratab, S. Agarwal, and T. Meyarivan. 2002. A Fast and Elitist Multi-Objective Genetic Algorithm. NSGA-II, IEEE Trans. Evol. Computation, 6(2), pp. 182-197.

Herrera, G.S., and G.F. Pinder. 1998. Cost-Effective Groundwater Quality Sampling Network Design. Proceedings XII International Conference on Computational Methods in Water Resources, Vol. I, Crete, Greece, June.

Johnson, V.M., M.N. Ridley, R.C. Tuckfield, and R.A. Anderson. 1996. Reducing the Sampling Frequency of Ground Water Monitoring Wells. Environmental Science and Technology, 30, 355-358.

Nobel, C. 2003. Applying a Three-Tiered Approach to Monitoring Network. Seventh International Battelle In Situ and On-Site Bioremediation Symposium, June.

OMB. 2008. Discount Rates for Cost-Effectiveness, Lease Purchase, and Related Analyses. OMB Circular No. A 94, Appendix C (www.whitehouse.gov/omb/circulars/a094/a94_appx-c.html).

Reed, P, B.S. Minsker, and D.E. Goldberg. 2003. Simplifying Multiobjective Optimization: An Automated Design Methodology for the Nondominated Sorted Genetic Algorithm-II. Water Resources Research 39(7):1196.

Rizzo, D.M., D.E. Dougherty, and M.Yu. 2000. An Adaptive Monitoring and Operations System (aLTMOsTM) for Environmental Management. In Building Partnerships, Proceeding. 2000 ASCE Joint Conference on Water Resources Engineering and Water Resources Planning & Management, R. H. Hotchkiss and M. Glade (eds.), American Society of Civil Engineers, Reston, VA, CD-ROM (ISBN 0-7844-0517-4).

Summit Envirosolutions. 2008. Sampling Optimizer and Data Tracker: Reference Manual.

This page left blank intentionally.

APPENDIX A

POINTS OF CONTACT

Point of Contact	Organization	Phone Fax E-Mail	Role
Ms. Karla Harre	U.S. Navy NAVFAC ESC Code 414, 1100 23rd Avenue Port Hueneme, CA 93043	(805) 982-2636 (805) 982-4304 karla.harre@navy.mil	Principal Investigator
Mr. Tanwir Chaudhry	U.S. Navy NAVFAC ESC Code 414, 1100 23rd Avenue Port Hueneme, CA 93043	(805) 982-1609 (805) 982-4304 tanwir.chaudhry@navy.mil	Co-Principal Investigator
Mr. Robert Greenwald	GeoTrans, Inc. 2 Paragon Way Freehold, NJ 07728	(732) 409-0344 (732) 409-3020 rgreenwald@geotransinc.com	Technical Consultant, Contractor
Mr. Matthew Zavislak	Summit Envirosolutions, Inc. 1217 Bandana Boulevard North St. Paul, MN 55108	(651) 842-4248 (651) 647-0888 mzavislak@summite.com	Software Developer
Dr. Barbara Minsker	Independent Consultant 2508 Waterbury Place Champaign, IL 61822	(217) 417-4198 (217)-333-4050 barbara.minsker@gmail.com	Technical Consultant, Contractor
Dr. Charles Davis	EnviroStat 3468 Misty Court Las Vegas, NV, 89120	(702) 456-8994 (702) 456-8114 charles.davis@envirostat-nv.com	Technical Consultant, Contractor
Dr. Andrea Leeson	ESTCP Office 901 N. Stuart Street Suite 303 Arlington, VA 22203	(703)696-2118 (703)696-2114 andrea.leeson@osd.mil	Environmental Restoration Program Manager

ESTCP Program Office

901 North Stuart Street
Suite 303
Arlington, Virginia 22203

(703) 696-2117 (Phone)
(703) 696-2114 (Fax)

E-mail: estcp@estcp.org
www.estcp.org